IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Re Application of: Rämö et al.	:	Attorney Docket No.: 944-003.182
Serial No.: 10/692,290	:	Examiner: Michael N. Opsasnick
Filed: October 23, 2003	:	Art Unit: 2626

For: METHOD AND SYSTEM FOR SPEECH CODING

Mail Stop Appeal Brief - Patents
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

BRIEF OF APPELLANTS (37 CFR §41.37)

Sir:

This is an appeal from the final rejection contained in a Final Office Action mailed on September 8, 2008, (the "Final Office Action"), rejecting claims 1 and 3-41 and 49-56. A notice of Appeal was filed on November 2, 2010. This Appeal Brief, as a replacement of the Appeal Brief filed on January 3, 2011, is filed in response to Notification of Non-Compliant Appeal Brief, mailed January 25, 2011.

****If any fee and/or extension is required in addition to any enclosed herewith, please charge Account No. 23-0442.***

CERTIFICATE OF MAILING/TRANSMISSION (37 C.F.R. § 1.8(a))

I hereby certify that this correspondence is, on the date shown below, being:

MAILING

☒ Deposited with the United States Postal Service with sufficient postage as first class Mail in an envelope addressed to the Commissioner for Patents, PO Box 1450, Alexandria, VA 22313-1450.

FACSIMILE

☐ Transmitted by facsimile to the U.S. Patent and Trademark Office.

Date: February 24, 2011

Marie E. Forte
Signature
Marie E. Forte
(type or print name of person certifying)

I. REAL PARTY IN INTEREST (37 CFR §41.37(c)(1)(i))

The real party in interest in this action is Nokia Corporation, Keilalahdentie 4, FIN-02150 Espoo, Finland, by virtue of the Assignment dated November 10 and 14, 2003. The Assignment was recorded in the U.S. Patent and Trademark Office on February 9, 2004, Reel 014970 and Frame 0234

II. RELATED APPEALS AND INTERFERENCES (37 CFR §41.37(c)(1)(ii))

There are no related appeals or interferences.

III. STATUS OF CLAIMS (37 CFR§41.37(c)(1)(iii))

The status of the claims is:

Claims pending: 1, 3-41 and 49-56

Claims objected to: none.

Claims rejected: 1, 3-41 and 49-56

Claims on appeal: 1, 3-41 and 49-56

IV. STATUS OF AMENDMENTS (37 CFR §41.37(c)(1)(iv))

No amendment of claims 1, 3-41 and 49-56 has been filed subsequent to final rejection.

V. SUMMARY OF CLAIMED SUBJECT MATTER (37 CFR §41.37(c)(1)(v))

Appellant's invention is directed to a method and device related to the segmentation of an audio signal into a plurality of segments and the encoding of the segments with different encoding settings. The segmentation is chosen such that the intra-segment similarity of the speech parameters is high (page 11, lines 19-20). The segmentation can be made based on quantized or unquantized parameters (page 13, lines 4-5). After the segmentation, the segments can be classified into types so that each segment can be coded by a coding scheme based on the segment type (page 11, lines 18-25). In particular, the characteristics of the audio signal are indicated in the speech parameters extracted from a parameter extraction unit 12 (page 13, line 9-11), and the partitioning or segmentation is carried out by a compression module 20 based on the behavior of the parameters (page 13, lines 21-24).

The speech parameters are extracted at regular intervals including linear prediction coefficients, speech energy or gain, pitch and voicing information (page 11, lines 26-27). The pitch associated with the speech signal is shown in Figure 2b, the voicing information associated with the speech signal is shown in Figure 2c and the energy associated with the speech signal is shown in Figure 2d. The claimed invention uses those audio characteristics for partitioning the audio signal into a plurality of segments. For example, a segmentation algorithm can be implemented based on a number of audio characteristics (page 12, lines 1-24). An example of the audio signal segmentation, according to present invention, is shown in Figures 3a to 3d. Figure 3a shows an audio signal from frames 100 to frames 200. The energy associated with that audio signal is shown in Figure 3b and the voicing information associated with that audio signal is shown in Figure 3c. Based on the energy and the voicing information, the audio signal is segmented into 7 segments as shown in Figure 3d. Because the segments of the audio signal based on the audio characteristics will likely have different parameters associated with the audio characteristics, each segment can be efficiently coded using a coding scheme in order to meet the perceptual requirements, for example (page 12, lines 25-29). Thus, according to the claimed method, the partitioning of the audio signal is carried out based on the parameters indicative of the audio characteristics of the audio signal, and the segments are encoded with different encoding settings.

The invention of independent claim 1 is directed to a method for partitioning an audio signal into a plurality of segments based on parameters indicative of audio characteristics of the audio signal (page 12, lines 1-8; Figures 3a-3d; page 12, line 29 – page 13, line 7), the parameters are obtained from an audio signal for each of a plurality of consecutive time intervals (page 11, lines 26-27), and encoding the segments with different encoding settings (page 11, lines 24-25; page 12, lines 25-28).

The invention of independent claim 19 is directed to a decoder (item 40, Figure 4). The decoder comprises an input for receiving audio data and a module for generating a further audio signal (page 13, lines 15-20). The audio data is indicative of a plurality of segments obtained by partitioning the audio signal based on parameters indicative of audio characteristics of the audio signal, and extracted from each of a plurality of consecutive time intervals (Figures 3a-3d; page

12, line 29 – page 13, line 7; page 13, lines 9-11; page 11, lines 26-27). The audio data is also indicative of an adjusted representation of the parameters so that the further audio signal is generated based on the adjusted representation and the encoding settings (page 21, lines 21-26).

The invention of independent claim 22 is directed to an encoding device (item 20, Figure 4). The encoding device comprises an input for receiving audio data indicative of parameters (item 112, Figure 4), and an adjustment module for adjusting one or more of the parameters for providing an adjusted representation of the parameters, wherein the adjustment comprises partitioning the audio signal into a plurality of segments based on the parameters obtained for the consecutive time intervals and encoding the segments based on one or more of a plurality of encoding settings (page 13, lines 8-12, lines 21-28; page 21, lines 21-26).

The invention of independent claim 27 is directed to an electronic device (item 40, Figure 4). The device comprises:

an input module for receiving audio data indicative of a plurality of segments of an audio signal (item 120, Figure 4), wherein one or more parameters are extracted from the audio signal for each of a plurality of consecutive time intervals, the parameters indicative of audio characteristics of the audio signal (page 11, lines 26-27), and wherein the plurality of segments are obtained by partitioning the audio signal based on the parameters extracted for the consecutive time intervals (page 12, lines 1-8; Figures 3a-3d; page 12, line 29 – page 13, line 7; page 13, lines 9-11), and the audio data is indicative of the parameters in an adjusted representation (page 21, lines 21-32); and

a decoder, responsive to the audio data, for generating a synthesized audio signal based on the adjusted representation (page 13, lines 15-20).

The invention of independent claim 31 is directed to a communication network (Figure 11). The network comprises a plurality of base stations; and a plurality of mobile stations adapted for communicating with the base stations (page 23, line 26-31), wherein at least one of the mobile stations (item 50, Figure 4) comprises:

an input module for receiving audio data from at least one of the base stations, the audio data indicative of a plurality of segments of an input audio signal (item 120, Figure

4) , wherein one or more parameters are extracted from the audio signal for each of a plurality of consecutive time intervals, the parameters indicative of audio characteristics of the audio signal (page 11, lines 26-27), and wherein the plurality of segments are obtained by partitioning the input audio signal based on the parameters extracted for the consecutive time intervals (page 12, lines 1-8; Figures 3a-3d; page 12, line 29 – page 13, line 7; page 13, lines 9-11), and encoded with a plurality of encoding settings based on the audio characteristics (page 12, lines 25-28), the audio data indicative of the parameters in an adjusted representation (page 21, lines 21-32).

a decoder, responsive to the audio data, for generating a synthesized audio signal based on the adjusted representation (page 13, lines 15-20).

In the invention of dependent claim 3, the audio characteristics include voicing characteristics in the segments of the audio signal (Figure 2c).

In the invention of dependent claim 4, the audio characteristics energy characteristics in the segments of the audio signal (Figure 2d).

In the invention of dependent claim 5, the audio characteristics include pitch characteristics in the segments of the audio signal (Figure 2b).

In the invention of dependent claim 6, the partitioning of the audio signal into segments is carried out concurrent to the encoding of the segments. As disclosed, a quantization mode is selected for each segmented parameter signal with k parameter values within the segment, but a reduced number i of parameters values are coded by the quantizer into the bitstream (page 15, lines 7-24).

In the invention of dependent claim 7, the partitioning is carried out before said encoding (page 7, lines 20-21).

In the invention of dependent claim 8, a plurality of voicing values are assigned to the audio characteristics of the audio signal in said segments, and the partitioning is carried out based on the assigned voicing values (page 11, lines 22-25).

In the invention of dependent claim 9, the plurality of values includes a value designated to a voiced speech signal and another value designated to an unvoiced signal (paragraph [0100]).

In the invention of dependent claim 10, the plurality of values further includes a value designated to a transitional stage between the voice and unvoiced signal (page 11, lines 28-31).

In the invention of dependent claim 11, the plurality of values further includes a value designated to an inactive period in the audio signal (page 11, lines 22-25).

In the invention of dependent claim 12, the encoding includes selecting a quantization mode for improving bit allocation and for reducing_parameter update rate, and the partitioning is carried out based on the selected quantization mode (page 14, lines 27-31).

In the invention of dependent claim 13, the partitioning is carried out based on a selected target accuracy in reconstructing of the audio signal, wherein the target accuracy is selected based on a distortion criteria comparing upsampled quantized values and modified parameter signal (page 14, lines 30-35; page 16, lines 22-32).

In the invention of dependent claim 14, the partitioning also includes providing a linear pitch representation in at least some of the segments (page 21, lines 21-28).

In the invention of dependent claim 15, the audio signal is encoded into audio signal data, and the method further comprises: forming a parameter signal based on the audio signal data having a first number of signal data; downsampling the parameter signal to a second number of signal data for providing a further parameter signal, wherein the second number is smaller than the first number; and upsampling the further parameter signal to a third number of signal data in decoding, wherein the third number is greater than the second number (page 15, lines 13-33).

In the invention of dependent claim 16, the third number is equal to the first number (page 15, lines 25-27).

In the invention of dependent claim 17, the signal data comprises quantized parameters (page 13, lines 4-5).

In the invention of dependent claim 18, the signal data comprises unquantized parameters (page 13, lines 4-5).

In the invention of dependent claims 33, 37, 38, 39, 40, the encoding settings comprise bit allocation, quantization accuracy, quantization method and parameter update rate (Table II; page 19, lines 12-24; page 20, lines 16-20).

In the invention of dependent claim 34, the audio signal contains sinusoidal components and said parameters include frequency values, amplitude values and phase values indicative of the sinusoidal components (page 2, line 27 – page 3, line 11).

In the invention of dependent claim 35, the parameters include pitch, voicing, amplitude and energy of the audio signal (Figure 2).

In the invention of dependent claim 36, the parameters include pitch contour data containing a plurality of pitch values representative of an audio segment in time (Figure 10; page 21, lines 25-26).

In the invention of dependent claim 41, the audio signal comprises a plurality of frames and the audio signal in each frame has a waveform and wherein a further audio signal is produced in the decoding stage independently of the waveform (page 13, lines 8-20).

In the invention of dependent claims 49, 51, 53, the parameters are obtained from the audio signals in regular time intervals (page 11, lines 26-27).

In the invention of dependent claims 50, 52, 54, 55, 56 the partitioning is based on the similarity in the parameters among consecutive time intervals (page 11, lines 23-24).

The dependent claim 26 is directed to a computer readable storage medium embedded with a computer program having programming code for carrying out the method of claim 1 (Figure 4; page 13, lines 17-20).

In the invention of dependent claim 20, 28, the audio data is recorded on an electronic medium, and wherein input of the decoder is operatively connected to the electronic medium for receiving the audio data (Figure 4; page 13, lines 15-17).

In the invention of dependent claim 21, 29, the audio data is transmitted through a communication channel, and wherein the input of the decoder is operatively connected to the communication channel for receiving the audio data (page 13, lines 15-17).

In the invention of dependent claim 32, the parameters including pitch contour data containing a plurality of pitch values representative of an audio segment in time, and wherein the pitch contour data in the audio segment in time is approximated by a plurality of consecutive sub-segments in the audio segment for providing a plurality of end points, and wherein the end points include a first end point and a second end point for defining each of said sub-segments; and the decoder also includes a reconstruction module for reconstructing the audio segment based on the received audio data (Figure 10).

In the invention of dependent claim 23, the encoding device also comprises a quantization module, responsive to the adjusted representation, for coding the parameters in the adjusted representation (page 13, lines 15-20).

In the invention of dependent claim 24, the encoding device also comprises an output end, operatively connected to a storage medium, for providing data indicative of the coded

parameters in the adjusted representation to the storage medium for storage (Figure 4; page 13, lines 15-17).

In the invention of dependent claim 25, the encoding device also comprises an output end, operatively connected to a communication channel, for providing signals indicative of the coded parameters in the adjusted representation to the communication channel for transmission (Figure 4; page 13, lines 15-17).

In the invention of dependent claim 30, the electronic device comprises a mobile terminal (page 13, lines 15-17).

VI. GROUNDS OF REJECTION TO BE REVIEWED ON APPEAL (37 CFR §41.37(c)(1)(vi))

At section 2 of the final office action, claims 1, 3-42, 49-46 are rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention. The Examiner states that in claims 1, 19, 22, 27, 31 and 32, the claim recitations pertaining obtaining/segmenting, for a plurality of consecutive time intervals, audio signals based on audio characteristics are vague and indefinite because it is not clear as to which segmenting aspect of the disclosure this refers. The Examiner further states that the specification discloses two aspects of segmenting:

1) a typical audio encoder that extracts audio signal information (outputting segments based upon voice/unvoiced, silence decision denoted as line 110 into the sub-block 12 in Figure 4, generating segmented audio with associated parameters 112 (page 13, lines 8-14 of the specification), and

2) the sub-block 20 re-segments sequence of initial segments based on degree of voicing, etc., derived from speech parameters (Figure 4, page 15, lines 1-17 of the specification).

The Examiner further states that the current claim scope does not distinguish between these two sections of the applicant's disclosure and as such, these claims are rejected under 35 U.S.C. 112, second paragraph. For art related examination purposes only, the Examiner will interpret the claim scope to read upon the first section (aspect) discussed above, namely, the encoder of Figure 4 that encompasses only line 110, sub-block 12, and line 112. The dependent claims do not remedy the deficiencies of the independent claims, and as such, are also rejected under 35 U.S.C. 112, second paragraph.

At section 3, claims 1, 3-14, 19-21, 26-37, 39-44 and 46-48 are rejected under 102(b) as being anticipated by *Gersho et al.* (U.S. Patent No. 6,311,154, hereafter referred to as *Gersho*).

In particular, in rejecting independent claim 1, the Examiner states that *Gersho* discloses segmenting {partitioning or classifying} the audio input signal {speech} into a plurality of segments {frames} (partitioning samples of speech signal into frames, col.4,lines 25-27) based on the audio characteristics {classes} of the audio signal (classifying the speech signal in each frame into one of a plurality of classes, col.4, lines 25-27); and encoding the segments {frames} with different encoding settings {excitation} (encoding an excitation for the frame using one of

the plurality of excitation coding ... selected according to the class of the frame, col.4, lines 30-33).

In rejecting independent claims 19 and 27 under 102(b), the Examiner states that *Gersho* discloses an input for receiving audio data indicative of parameters in the adjusted representation (Figure 3, input applied to filter 14), and a module for generating the audio signal based on the adjusted representation (Figure 3). The Examiner states that it would have been inherent to one skilled in the art to use a decoder to reverse the encoding data for further processing, such as modulating or storing the audio signal.

In rejecting independent claim 31, the Examiner states that *Gersho* discloses a cell phone system having both a base station and a mobile station (col.6, lines 33-36); a decoder (Figures 1, 4, 5, 9; col.8, lines 54-63); and an input for receiving audio data (Figures 1, 4, 5, col.3, lines 1-15).

At section 5, claims 15-18, 22-25, 38 and 45 are rejected under 102(e) as being anticipated by *Sinha et al.* (U.S. Patent No. 7,191,136 B2, hereafter referred to as *Sinha*).

In rejecting claims 15, 22, 23 and 45, the Examiner states that *Sinha* discloses a method for use in parametric audio coding to encode an audio signal by segmenting the audio signal, for each of a plurality of consecutive time intervals, one or more parameters from an audio signal, the one or more parameters relating to audio characteristics of the audio based on audio characteristics of the audio signal (by high-pass filtering the input audio signal as disclosed in col.4, lines 47-59) and then performing a non-linear parametric representation of the signal (col. 4, lines 53-59)).

VII. ARGUMENT (37 CFR §41.37(c)(1)(vii))

A. 112 Rejection

As pointed out in Subsections A.1 to A.2 below, the Examiner **errs** in interpreting Figure 4 of the present invention.

In the rejection of claims 1, 3-42, 49-46 under 35 U.S.C. 112, second paragraph, the Examiner states that claims 1, 19, 22, 27, 31 and 32 have the limitation of segmenting audio signals based upon audio characteristics, but it is not clear as to which segmenting aspect of the disclosure this refers. The Examiner states that the specification discloses two aspects of segmenting:

1) a typical audio encoder that extracts audio signal information (outputting segments based upon voice/unvoiced, silence decision as shown in line 110 into the sub-block 12 in Figure 4, generating segmented audio with associated parameters 112 (page 13, lines 8-14 of the specification), and

2) the sub-block 20 re-segments the sequence of initial segments based on the degree of voicing, etc., derived from speech parameters (Figure 4, page 15, lines 1-17 of the specification).

The Examiner further states that the current claim scope does not distinguish between these two sections of the applicant's disclosure and as such, these claims are rejected under 35 U.S.C. 112, second paragraph.

A.1 Examiner Errs in Interpreting Figure 4

The Examiner errs in interpreting the speech coding system as shown in Figure 4 and the description thereof. In particular, the Examiner errs in interpreting what block 12 and line 112 are. The Examiner further errs in stating the functions of a typical audio encoder.

A.1.1. Block 12 in Figure 4

Block 12 is labeled as an encoder in Figure 4. Its function is to extract parameters from the input signal 110 (page 13, lines 9-11). A typical parametric speech coder is used to estimate parameters at regular intervals (page 5, lines 22-25). There is no segmentation involved. So far as the encoder 12 is concerned, there is no outputting segments based upon voice/unvoiced, silence decision. *See* Subsection D.2 below.

A.1.2 Line 112 in Figure 4

Line 112 in Figure 4 is labeled as parameters which are the only output from the encoder 12. Line 112 represents speech parameters (page 13, lines 8-9). Line 112 does not contain segments that are outputted based on voiced or unvoiced. *See* Subsection D.3 below.

A.1.3 Block 20 in Figure 4

As disclosed, block 20 is described as compression module, which is used to segment the input speech signal based on the behavior of the parameters (page 13, line 21-24). The parameters are indicated in line 112. The Examiner errs in stating that block 20 is used to output the re-segmented sequence of initial segments based on a degree of voicing.

A.2 Figure 4 Depicts only One Segmentation Block

In Figure 4, block 12 is used for extracting parameters from audio signal 110, and block 20 is used to partition the audio signal into segments based on the extracted parameters 112. The specification does not disclose two aspects of segmenting as alleged by the Examiner. The 112 rejection is improper.

B. 102 Rejection over *Gersho*

At issue here is whether *Gersho* discloses partitioning the audio signal into frames based on classes or audio characteristics of the audio signal.

As pointed out in Subsection E below, *Gersho* classifies each of the frames into classes only after partitioning the speech signal into frames. Therefore, *Gersho* does not disclose partitioning the audio signal into frames based on classes of the audio signal.

B.1 The Examiner Errs in Interpreting *Gersho*

Examiner **errs** in equating "segmenting" to "classifying".

On page 3 of the office action, the Examiner states that *Gersho* discloses segmenting {partitioning or classifying} the audio input signal into a plurality of segments {frames} (partitioning samples of speech signal into frames), and obtaining, for each of a plurality of consecutive time intervals, one or more parameters from an audio signal, said one or more parameters relating to audio characteristics {classes} of the audio signal. The Examiner points to

col.4, lines 25-27 to show that *Gersho* discloses classifying the frames in the speech signal into one of the plurality of classes.

The Examiner is correct in stating that *Gersho* discloses classifying the frames in the speech signal into one of the plurality classes. This means that *Gersho* classifies each of the frames into classes after segmenting or partitioning the speech signal into frames (see Subsection E below).

However, the Examiner **errs** in equating "segmenting" to "classifying".

It is respectfully submitted that classifying and partitioning are different processes. It is improper for the Examiner to equate "segmenting" to "classifying" without citing any references.

B.2 Claimed Invention

Claim 1 includes the limitations of

1) obtaining, for each of a plurality of consecutive time intervals, one or more parameters from an audio signal, said one or more parameters indicative of audio characteristics of the audio signal, and

2) partitioning the audio signal into a plurality of segments based on the parameters obtained for the consecutive time intervals.

B.2.1 *Gersho* Fails to Anticipate Claim 1

In the claimed invention, the partitioning of the audio signal into segments is based on the parameters indicative of the audio characteristics of the audio signal. In *Gersho*, each of the frames is classified into classes only after segmenting or partitioning the speech signal into frames. Since the class information is not available when the partitioning is carried out, *Gersho* does not disclose partitioning the audio signal into segments based on the parameters indicative of the audio characteristics of the audio signal.

For the above reasons, *Gersho* fails to anticipate claim 1.

B.2.2 *Gersho* Fails to Anticipate Claims 19, 27 and 31

Claims 19, 27 and 31 include the limitation that the plurality of segments are obtained by partitioning the audio signal based on parameters indicative of audio characteristics of the audio signal.

In rejecting claims 19 and 27, the Examiner states that *Gersho* discloses an input for receiving audio data indicative of the parameters, but fails to point out that where *Gersho* discloses that the segments are obtain by partitioning the audio signals based on the parameters.

In rejecting claim 31, the Examiner broadly refers to Figures 1, 4-5, and 9, Abstract and col.8, lines 54-63, but fails to specifically point out where *Gersho* discloses "said adjusting comprises the steps of segmenting the audio signal into a plurality of segments based on the characteristics of the audio signal".

As pointed out in Sub-section B.2.1 above, *Gersho* does not disclose partitioning the audio signal into segments based on the parameters indicative of the audio characteristics of the audio signal.

For the above reasons, *Gersho* fails to anticipate claims 9, 27 and 31.

B.2.3 *Gersho* Fails to Anticipate Claims 3-14, 20, 21, 26, 28-30, 32-37, 39-41, 49-56

Dependent claims 3-14, 20, 21, 26, 28-30, 32-37, 39-41, 49-56 are dependent from claims 1, 19, 27 and 31 and include further limitations. For reasons regarding claims 1, 19, 27 and 31 as pointed out in Subsections B.2.1 and B.2.2 above, *Gersho* also fails to anticipate claims 3-14, 20, 21, 26, 28-30, 32-37, 39-41, 49-56.

C. 102 Rejection over *Sinha*

The 102 rejection over *Sinha* is improper.

In rejecting independent claim 22, and dependent claims 15, 23 and 45, the Examiner states that *Sinha* discloses segmenting the audio signal for each of the plurality of consecutive time intervals, one or more parameters from an audio signal, the parameters relating to audio characteristics of the audio based on audio characteristics of the audio signal (by high pass filtering the input audio signal as disclosed in col.4 lines 47-51).

The claimed invention **has nothing to do** with segmenting an audio signal for each of the plurality of consecutive time intervals into one or more parameters. The claimed invention is concerned with partitioning the audio signal into a plurality of segments based on the parameters obtained from the audio signal.

C.1 *Sinha* Fails to Anticipate Claim 22

Independent claim 22 includes the limitation of partitioning the audio signal into a plurality of segments based on the parameters obtained in a plurality of consecutive time intervals, the parameters indicative of audio characteristics of the audio signals.

The Examiner fails to point out where *Sinha* discloses obtaining parameters from an audio signal and partitioning the audio signal into a plurality of segments based on the parameters. As known in the art, high pass-filtering is not the same as 1) obtaining parameters in a plurality of consecutive time intervals and 2) partitioning the audio signal into a plurality of segments based on the parameters.

Thus, *Sinha* fails to anticipate claim 22.

C.2 *Sinha* Fails to Anticipate Claim 23-25 and 38

Claims 23-25 and 38 are dependent from claim 22 and include further limitations. For reasons regarding claim 22 as pointed out in Subsection C.1 above, *Sinha* also fails to anticipate claims 23-25.

C.3 *Sinha* Fails to Anticipate Claims 15-18

Claims 15-18 are dependent from claim 1. Claim 1 includes the limitation of: obtaining, for each of a plurality of consecutive time intervals, one or more parameters from an audio signal, said one or more parameters indicative of audio characteristics of the audio signal, and partitioning the audio signal into a plurality of segments based on the parameters obtained for the consecutive time intervals.

Sinha does not disclose the limitation of claim 1.

Thus, *Sinha* fails to anticipate claims 15-18.

D. General Description of Figure 4

On page 13, lines 8-14 of the specification, it is disclosed that

Figure 4 is a speech coding system that quantizes speech parameters 112 utilizing the segmentation information. The compression module 20 can use either quantized parameters from an existing speech coder, or the compression module 20 can use the unquantized parameters

directly coming from the parameter extraction unit 12. Moreover, a pre-processing stage (not shown) may be added to the encoder to generate speech signals with specific energy level and/or frequency characteristics. The input speech signal 110 can be generated by a human speaker or by a high-quality TTS algorithm.

D.1 Line 110 in Figure 4

As shown in Figure 4, line 110 represents an input signal before it is conveyed to the encoder 12. The input signal 110 can be generated by a human speaker or by a high-quality text-to-speech (TTS) algorithm. The input signal 110 corresponds to the "audio signal" as claimed. As this point, the input signal is not segmented by any device.

D.2. Block 12 in Figure 4

Block 12 is labeled as an encoder as shown in Figure 4. Its function is to extract parameters for the input signal 110 (page 13, lines 9-11). Therefore, the encoder is also known as a parametric speech coder as described on page 5, lines 22-25 as follows:

In a typical parametric speech coder, the speech parameters are estimated from the speech signal at regular intervals. The length of this interval is usually equal to the frame length used in the coder. While some parameters (e.g. pitch) may be estimated more often than others, the estimation rate for a given parameter is usually constant.

On page 11, lines 26-31, it is disclosed that:

In a typical parametric speech coder, the parameters extracted at regular intervals include linear prediction coefficients, speech energy (gain), pitch and voicing information. To illustrate the speech signal segmentation method of the present invention, it is assumed that the voicing information is given as an integer value ranging from 0 (completely unvoiced) to 7 (completely voiced), and that the parameters are extracted at 10 ms intervals.

The speech coder used in the claimed invention is for parameter extraction. There is no segmentation involved. Segmentation means sectioning, partitioning or dividing. There is no

indication in the disclosure that the parameter extraction unit 12 partitions or divides the input speech signal into separated segments through the parameter extraction process, even though parameter extraction can be carried out in regular intervals (page 11, lines 26-32). So far as the encoder 12 is concerned, there is no outputting segments based upon voice/unvoiced, silence decision.

D.3. Line 112 in Figure 4

Line 112 in Figure 4 is labeled as parameters which are the only output from the encoder 12. On page 13, lines 8-9, it is disclosed that

Figure 4 is a speech coding system that quantizes speech parameters 112 utilizing the segmentation information.

Before the speech parameters 112 are conveyed to and processed by the compression module 20, the speech parameters 112 are not quantized in a coding process. According to the present invention, quantization of the speech parameters is based on the segmentation information. Segmentation information is obtained in the compression module 20 as "segmented parameter signal", which is formed from a plurality of parameter values inside a segment (page 14, lines 27-29; page 15, lines 12-14).

Thus, according to the present invention, segmentation has not been carried out before the speech parameters 112 are processed by the compression module. Line 112 represents speech parameters. Line 112 does not contain voiced speech segments or unvoiced speech segments.

This is to show that the so-called first aspect of segmenting as asserted by the Examiner is wrong.

E. The Cited *Gersho* Reference

According to *Gersho*, for the purpose of performing linear-predictive (LP) analysis on the input speech, and for the purpose of packaging the data to be transmitted into a fixed number of bits for each fixed frame interval, the speech encoder has a fixed (basic) frame structure. Each basic frame is partitioned or segmented into M equal or nearly equal length basic subframes (col. 7, lines 18-26; Figure 2). According to *Gersho*, in conventional analysis-by-synthesis (AbS)

coding schemes, the excitation signal for each subframe is selected by a search operation. It is difficult or impossible to obtain an adequately precise representation of the excitation segment using the conventional schemes (page 7, lines 27-33).

Gersho sets out to improve the AbS coding method by locating the actual time location of the active intervals in a sub-frame so that the coding effort can be concentrated with the windows corresponding to the active intervals. Active intervals are certain naturally-occurring intervals of the excitation signal which contain most of the important activity (col.7, lines 34-50). *Gersho* adaptively modifies the sub-frame boundaries and determines the window sizes and locations within sub-frames (col.2, lines 46-50). *Gersho* uses a pattern classifier to determine a classification that best describes the character of the speech signal in each frame (col.2, line 56-64). The method for coding a speech signal, according to *Gersho*, includes: **1) partitioning samples of a speech signal into frames**; 2) deriving a residual signal for each frame; **3) classifying the speech signal in each frame into one of a plurality of classes**; 4) identifying the location of at least one window in the frame by examining the residual signal for the frames; and 5) encoding the excitation for the frame based on the class of the frame (col.4, lines 23-34).

Since classification in **step 3** above is carried out after the partitioning in **step 1**, class information is not available before the speech signal is segmented or partitioned into frames. Even in the conventional AbS schemes, excitation is searched after the speech signal is segmented into frames and into sub-frames. *Gersho* does not disclose segmenting the input speech signal into segments based on classes.

F. The Cited *Sinha* Reference

Sinha is concerned with a coding scheme which compresses information consisting of coded low frequency components as well as parametric representations for the high frequency components from the high pass filter (Abstract, column 4, lines 44-49). In particular, *Sinha* allows the input signal to pass through both a high pass filter and a low-pass filter so that the audio components in the high-frequency range and the audio components in the low-frequency range are encoded using different models. While the audio components can be encoded with parameters in a parametric representation and the audio characteristics of audio components can be indicative of parameters in the parametric representation, high frequency range or low frequency range is not a parameter in the parametric representations. Parameters, such as linear

prediction coefficients, speech energy (gain), pitch and voicing information, can be used for audio signal synthesis. *Sinha* does not disclose or suggest that the input audio signal is segmented based on audio characteristics indicative of parameters in a parametric representation.

VIII CLAIMS APPENDIX (37 CFR §41.37(c)(1)(viii))

1. A method, comprising:

obtaining, for each of a plurality of consecutive time intervals, one or more parameters from an audio signal, said one or more parameters indicative of audio characteristics of the audio signal,

partitioning the audio signal into a plurality of segments based on the parameters obtained for the consecutive time intervals; and

encoding the segments with different encoding settings.

2. (canceled)

3. A method according to claim 1, wherein the characteristics include voicing characteristics in said segments of the audio signal.

4. A method according to claim 1, wherein the characteristics include energy characteristics in said segments of the audio signal.

5. A method according to claim 1, wherein the characteristics include pitch characteristics in said segments of the audio signal.

6. A method according to claim 1, wherein said partitioning is carried out concurrent to said encoding.

7. A method according to claim 1, wherein said partitioning is carried out before said encoding.

8. A method according to claim 1, wherein a plurality of voicing values are assigned to the audio characteristics of the audio signal in said segments, and wherein said partitioning is carried out based on the assigned voicing values.

9. A method according to claim 8, wherein the plurality of values includes a value designated to a voiced speech signal and another value designated to an unvoiced signal.

10. A method according to claim 8, wherein the plurality of values further includes a value designated to a transitional stage between the voice and unvoiced signal.

11. A method according to claim 8, wherein the plurality of values further includes a value designated to an inactive period in the audio signal.

12. A method according to claim 1, wherein said encoding comprises selecting a quantization mode for improving bit allocation and for reducing_parameter update rate, wherein the partitioning is carried out based on the selected quantization mode.

13. A method according to claim 1, wherein said partitioning is carried out based on a selected target accuracy in reconstructing of the audio signal, wherein the target accuracy is selected based on a distortion criteria comparing upsampled quantized values and modified parameter signal.

14. A method according to claim 5, wherein said partitioning comprises providing a linear pitch representation in at least some of said segments.

15. A method according to claim 1, wherein the audio signal is encoded into audio signal data, said method further comprising:

forming a parameter signal based on the audio signal data having a first number of signal data;

downsampling the parameter signal to a second number of signal data for providing a further parameter signal, wherein the second number is smaller than the first number; and

upsampling the further parameter signal to a third number of signal data in decoding, wherein the third number is greater than the second number.

16. A method according to claim 15, wherein the third number is equal to the first number.

17. A method according to claim 15, wherein the signal data comprises quantized parameters.

18. A method according to claim 15, wherein the signal data comprises unquantized parameters.

19. A decoder, comprising:

an input for receiving audio data indicative of a plurality of segments of an audio signal, wherein one or more parameters are extracted from the audio signal for each of a plurality of consecutive time intervals, the parameters indicative of audio characteristics of the audio signal, and wherein the plurality of segments are obtained by partitioning the audio signal based on the parameters extracted for the consecutive time intervals, and the audio data is indicative of the parameters in an adjusted representation; and

a module, responsive to the audio data, for generating a further audio signal based on the adjusted representation and the encoding settings.

20. A decoder according to claim 19, wherein the audio data is recorded on an electronic medium, and wherein input of the decoder is operatively connected to the electronic medium for receiving the audio data.

21. A decoder according to claim 19, wherein the audio data is transmitted through a communication channel, and wherein the input of the decoder is operatively connected to the communication channel for receiving the audio data.

22. An encoding device comprising:

an input for receiving audio data indicative of parameters obtained from an audio signal in a plurality of consecutive time intervals, the parameters indicative of audio characteristics of the audio signal; and

an adjustment module for adjusting one or more of the parameters for providing an adjusted representation of the parameters, wherein said adjusting comprises partitioning the audio signal into a plurality of segments based on the parameters obtained for the consecutive

time intervals and encoding the segments based on one or more of a plurality of encoding settings.

23. An encoding device according to claim 22, further comprising a quantization module, responsive to the adjusted representation, for coding the parameters in the adjusted representation.

24. An encoding device according to claim 22, further comprising an output end, operatively connected to a storage medium, for providing data indicative of the coded parameters in the adjusted representation to the storage medium for storage.

25. An encoding device according to claim 22, further comprising an output end, operatively connected to a communication channel, for providing signals indicative of the coded parameters in the adjusted representation to the communication channel for transmission.

26. A computer readable storage medium embedded with a computer program comprising programming code for carrying out the method of claim 1.

27. An electronic device comprising:

an input module for receiving audio data indicative of a plurality of segments of an audio signal, wherein one or more parameters are extracted from the audio signal for each of a plurality of consecutive time intervals, the parameters indicative of audio characteristics of the audio signal, and wherein the plurality of segments are obtained by partitioning the audio signal based on the parameters extracted for the consecutive time intervals, and the audio data is indicative of the parameters in an adjusted representation; and

a decoder, responsive to the audio data, for generating a synthesized audio signal based on the adjusted representation.

28. An electronic device according to claim 27, wherein the audio data is recorded in an electronic medium, and wherein the input is operatively connected to the electronic medium for receiving the audio data.

29. An electronic device according to claim 27, wherein the audio data is conveyed through a communication channel, and wherein the input is operatively connected to the communication channel for receiving the audio data.

30. An electronic device according to claim 27, comprises a mobile terminal.

31. A communication network, comprising:

a plurality of base stations; and

a plurality of mobile stations adapted for communicating with the base stations, wherein at least one of the mobile stations comprises:

an input module for receiving audio data from at least one of the base stations, the audio data indicative of a plurality of segments of an input audio signal, wherein one or more parameters are extracted from the audio signal for each of a plurality of consecutive time intervals, the parameters indicative of audio characteristics of the audio signal, and wherein the plurality of segments are obtained by partitioning the input audio signal based on the parameters extracted for the consecutive time intervals and encoded with a plurality of encoding settings based on the audio characteristics, the audio data indicative of the parameters in an adjusted representation; and

a decoder, responsive to the audio data, for generating a synthesized audio signal based on the adjusted representation.

32. A decoder according to claim 19, the parameters including pitch contour data containing a plurality of pitch values representative of an audio segment in time, and wherein the pitch contour data in the audio segment in time is approximated by a plurality of consecutive sub-segments in the audio segment for providing a plurality of end points, and wherein the end points include a first end point and a second end point for defining each of said sub-segments; and

a reconstruction module for reconstructing the audio segment based on the received audio data.

33. A method according to claim 1, wherein the encoding settings comprise bit allocation, quantization accuracy, quantization method and parameter update rate.

34. A method according to claim 1, wherein the audio signal contains sinusoidal components and said parameters include frequency values, amplitude values and phase values indicative of the sinusoidal components.

35. A method according to claim 1, wherein the parameters include pitch, voicing, amplitude and energy of the audio signal.

36. A method according to claim 1, wherein the parameters include pitch contour data containing a plurality of pitch values representative of an audio segment in time.

37. A decoder according to claim 19, wherein the encoding settings include bit allocation, quantization accuracy, quantization method and parameter update rate.

38. An encoding device according to claim 22, wherein the encoding settings include bit allocation, quantization accuracy, quantization method and parameter update rate.

39. A computer readable storage medium according to claim 26, wherein the encoding settings include bit allocation, quantization accuracy, quantization method and parameter update rate.

40. A communication network according to claim 31, wherein the encoding settings include bit allocation, quantization accuracy, quantization method and parameter update rate.

41. A method according to claim 1, wherein the audio signal comprises a plurality of frames and the audio signal in each frame has a waveform and wherein a further audio signal is produced in the decoding stage independently of the waveform.

Claims 42-48. (canceled)

49. A method according to claim 1, wherein the parameters are obtained from the audio signals in regular time intervals.

50. A method according to claim 1, wherein said partitioning is based on the similarity in the parameters among consecutive time intervals.

51. A decoder according to claim 19, wherein the parameters are extracted from the audio signals in regular time intervals.

52. A decoder according to claim 19, wherein the plurality of segments are obtained based on similarity in the parameters among consecutive time intervals.

53. An encoding device according to claim 22, wherein the parameters are obtained from the audio signals in regular time intervals.

54. An encoding device according to claim 22, wherein said partitioning is based on similarity in the parameters among consecutive time intervals.

55. An electronic device according to claim 27, wherein the plurality of segments are obtained based on similarity in the parameters among consecutive time intervals.

56. A communication network according to claim 31, wherein said partitioning is based on similarity in the parameters among consecutive time intervals.

IX. EVIDENCE APPENDIX (37 CFR §41.37(c)(1)(ix))

There are no evidences submitted pursuant to 37 CFR §1.130, 1,131 or 1,132.

X. RELATED PROCEEDING APPENDIX (37 CFR §41.37(c)(1)(x))

There are no prior decisions rendered by a court or the Board in any proceeding identified pursuant to 37 CFR §41.37(c)(1)(ii).

CONCLUSION

It is respectfully submitted that the present invention as claimed is readily distinguishable over the cited *Gersho* and *Sinha* references. Appellants' invention is not disclosed in the applied prior art and there is no fair basis for alleging that appellants' invention is obvious in regard to such art.

In view of the above, it is respectfully submitted that the rejection of claims 1, 3-41 and 49-56 is in error and must be reversed.

Respectfully submitted,

Ken Lao

Date: February 24, 2011

Kenneth Q. Lao
Attorney for the Appellant
Registration No. 40,061

WARE, FRESSOLA, VAN DER SLUYS
& ADOLPHSON LLP
Bradford Green, Building Five
755 Main Street, P.O. Box 224
Monroe, CT 06468
Telephone: (203) 261-1234
Facsimile: (203) 261-5676
USPTO Customer No. 004955